The Origin of the Moon

Ye. L. Ruskol
O. Yu. Schmidt Institute of Earth Physics
Academy of Sciences,
Moscow, U.S.S.R.

This paper discusses fractionation of the chemical compositions of the Moon and the Earth and the thermal history of the Moon for formation of the Moon from an Earth-orbiting swarm of bodies, during the accumulation of the Earth.

This report will consider only a model for formation of the Moon in orbit around the Earth. This model is closely connected with the theory of formation of the Earth and the planets that has been developed in the O. Yu Schmidt Institute of Earth Physics, Academy of Sciences, Moscow (refs. 1 and 2). We will not dwell on a review of different hypotheses of lunar formation because they are given in published works (refs. 3, 4, and 5).

It is interesting to discuss those questions which traditionally are considered most difficult for this model: (1) differences in chemical composition of the Moon and the Earth and (2) the initial and boundary conditions of the thermal history of the Moon that conform to its current state.

Our physical-mechanical scheme of development for the Moon is built on the concept that during the active stage of growth of the Earth a satellite swarm of small bodies and particles formed around it. In order to become satellites of the Earth, these particles and bodies revolved in heliocentric orbits in the zone of the preplanetary cloud which the growing nucleus of the Earth was gradually sweeping out. As a consequence of inelastic collisions near the Earth, a certain fraction of the particles change to geocentric orbits. The satellite swarm of the Earth is very small. Its maximum radius is 100 times smaller than the distance from the Earth to the Sun, and that part of the swarm where an increased density over that of the "background" is achieved is tens of times smaller in size, such that the capture of particles in Earth orbit is most likely for near-Earth collisions. In this scheme, the Moon represents the final product of "assembly" of the particles and bodies of the satellite swarm.

If only the mutual collisions of particles in heliocentric orbits are taken into consideration, the mass of the satellite swarm proves to be several orders of magnitude less than the mass of the Moon. If the collision of free particles with those already captured is considered, an exponential growth of the mass of the swarm can be obtained and will overtake the mass growth of the Earth. However, if the density in some part of the cluster exceeds the "background" density by a factor of two or three orders of magnitude, the growth of an Earth satellite is activated in the swarm and quickly (in 10^2 to 10^3 yr) sweeps up newly entering particles, preventing too rapid growth of the swarm. With the process balanced in this manner, the maximum filling of the swarm is applied to the mass of the protoearth and is equal to two-fifths to one-half its modern mass (fig. 1). In order to obtain the required mass and geocentric moment of the swarm, it is not important how much time growth of the Earth takes because formation of the swarm is an accompanying process. However, chemical fractionation between the material of the

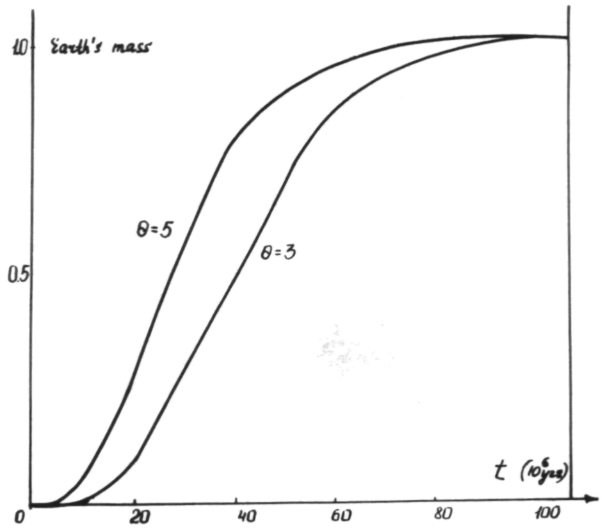

Figure 1.—*Rate of accretion of the Earth, according to Safronov (ref. 2).*

Earth and the material of the satellite swarm is not indifferent to the duration of the entire process. We assume a long time scale for growth of the Earth (10^8 yr), which is currently the most solidly based from the dynamical point of view (ref. 2).

Differences in Chemical Composition of the Earth and Moon During Formation of the Moon From the Earth Satellite Swarm

PARTIAL LOSS OF VOLATILE AND SEMIVOLATILE ELEMENTS AND ASSOCIATED ENRICHMENT OF THE SWARM WITH REFRACTORY SILICATES

Because the circumterrestrial swarm is secondary and formed later than the Earth, the material of the Moon should have been in a dispersed state longer: at first in the preplanetary cloud and then in circumterrestrial orbits. This lag in accumulation is at least 5×10^7 to 10^8 yr. During this time, the small bodies and particles experience numerous collisions at velocities determined by perturbations due to the protoearth: in the protoplanetary cloud the velocities are 3 to 4 km/s on the average and 5 to 7 km/s during capture into the swarm. A characteristic time for dissipation of light gases from the region of the terrestrial group of planets by the solar wind is 10^8 yr. As Kuiper showed (ref. 6), the solar wind is capable of ionizing and blowing the atoms of any element right out of the solar system if the space is free for radiation. It can be shown that the protoplanetary cloud was quite transparent in the radial direction from the Sun ($\tau \simeq 1$) out to the distance of the Earth from the Sun, if the radial distribution of the protoplanetary bodies followed the law $dN(a) \sim a^{-n_1} da$, where $n \leq 3.5$, with the mass of the material equal to the mass of the terrestrial planets and with $a_{max} \simeq 10^8$ cm. Much material evaporates in collisions of solid particles with velocities of several km/s; this shows up particularly in the low-melting and volatile components (H_2O, Pb, Bi, Tl, etc.). We (ref. 7) considered an example of selective removal of volatiles in a single collision, where the collision led to capture into the swarm. The colliding particles lose the evaporating components, a certain fraction recondenses into particles within the cluster, and the remaining ones are expelled from the swarm by the solar wind. Particles adsorbed by the growing Earth fall into it with high velocities (10 km/s) and also undergo evaporation. However, the strong gravitational field of the Earth prevents the evaporated material from escaping, and the inner, denser, and more opaque part of the swarm protects them from the action of the solar wind. Thus, two features should be considered as facilitating the Earth's acquiring relatively more volatiles and the Moon's acquiring relatively more refractory material: (1) earlier accumulation of the Earth from material that passed through a shorter sequence of collisions and (2) the presence of Earth gravity when the Earth and the satellite cluster simultaneously acquired material. It should be added that the work of Prof. E. Anders and colleagues (ref. 8), based on analyses of the first basalt samples returned from Mare Serenitatis in 1969 and in which the fact was established of depletion of volatiles from lunar samples and their enrichment in refractory elements, was an occasion for our discussion in 1971 of selective removal of volatiles from the particles of the circumterrestrial swarm. Subsequent investigations by the same authors and other groups showed how universal the regularity they discovered is for the Moon.

SELECTIVE REMOVAL OF THE SMALLEST PARTICLES BY CAPTURE IN THE SWARM AND ASSOCIATED ENRICHMENT OF THE LUNAR MATERIAL IN SILICATES

Let us proceed to the question of fractionation of iron and silicates. Studies of the size distribution of preplanetary bodies using coagulation theory, with allowance for fragmentation, have shown that this distribution asymptotically approaches the power law $dN(a) \sim a^{-n_1}\,da$, when $n_1 < 3 \cdot 5$ (for a mass distribution, this corresponds to the exponent 11/6) (refs. 2, 9, and 10). A simplifying assumption unavoidably has to be made in these studies, i.e., that the entire set of bodies and particles has uniform physical-mechanical properties, as well as composition. The predominant factor in the collisional interactions of the particles is mass, hence the important role of the largest bodies in the process of planetary growth. During the accretion of the terrestrial planets, the exponential-type distribution mentioned above becomes established among the protoplanetary bodies, where the size range is 12 to 13 orders of magnitude. At $n_1 \sim 3.5$ (and generally, at $3 < n_1 < 4$), there is an interesting feature that is illustrated in figure 2. A large part of the mass of the bodies is concentrated in the largest bodies, and a large part of the total surface is concentrated in the largest bodies, and a large part of the total surface is concentrated in the smaller fraction. This means that in process where the main role is played by the frequency of interactions, the small particles have the greatest activity. We have shown that in capture in the swarm the size distribution of the particles changes in the direction of enrichment with the smaller fraction (ref. 11). Thus, if the exponent for protoplanetary bodies was 3.5 after capture in the swarm (without considering the accumulating satellites) there would be a distribution with the exponent $n_2 \simeq 4.0$. Thus, there is a shift in the direction of smaller bodies. In this case, the mass is not concentrated in the largest bodies, but is uniformly distributed over the entire range of particle sizes.

Could this feature of the formation of the swarm lead to its enrichment in silicates or somehow to the depletion of iron? At the present time more data have appeared in favor of this possibility than there were in 1971 when we proposed this process based only on the qualitative considerations of Orowan (ref. 12). Recent experiments on collision of metallic particles have demonstrated that in a large velocity range, from 0.5 km/s to 10 km/s metallic particles weld themselves together, while at velocities less than 0.5 cm/s they undergo semielastic repulsion (ref. 13). Experiments with silicate particles confirmed the idea of Orowan that behavior of silicate particles is completely opposite to the behavior of metallic ones; i.e., in the impact velocity range from 1.5 km/s to 9.5 km/s destruction of the particles predominates over their agglomeration (ref.

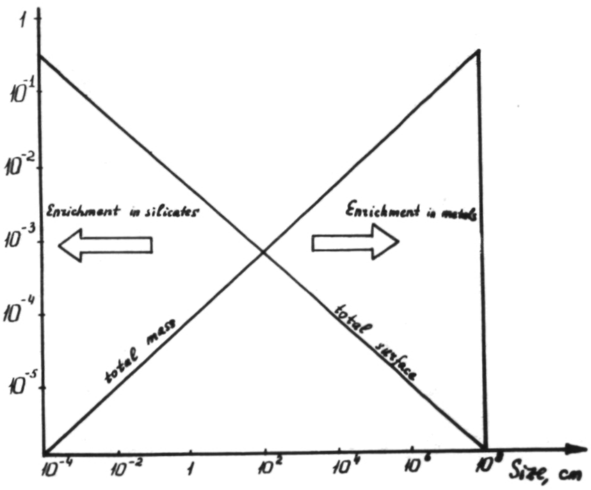

Figure 2.—*Characteristic features of the size distribution of bodies of the type* $dN(r) \sim r^{-3.5}\,dr$.

14). Only at velocities of less than 0.4 km/s is agglomeration possible (on condition that there is some adhesive agent on the surface of the silicate particles, e.g., frozen volatiles or an electrostatic charge). Investigations of the changes in materials "treated" by impacts with velocities of several km/s have shown that brittle material—silicates, hydrides—are broken up into very fine particles of micron and submicron sizes which cannot be obtained by other mechanical disintegration methods (ref. 15). But metals display only plastic deformation. The results of these experiments could be applied comparatively simply to protoplanetary particles if they were strictly divided into silicate and iron particles. However, a certain portion of the iron in the protoplanetary cloud could be oxidized and included in the particles in the form of FeO or Fe_3O_4. The oxides are rather more brittle than plastic materials. There are no similar experimental data for them in impacts with cosmic velocities. There is great interest in comparison of the results of impact treatment of iron oxides and silicon oxides as likely structural elements of the protoplanetary material in the region of the terrestrial group of planets. There are indications that at low velocities, on the order of 1 m/s, pulverization of silica (SiO_2) produces two times more total surface than pulverization of magnetite Fe_3O_4 (ref. 16). If the same behavior is displayed at cosmic velocities as a result of mutual collisions, a unique dependence of particle composition and size should develop. On the average, it could be represented in the form of a gradual increase in content of iron and metals in the most massive bodies and a gradual increase in silicate content in the smallest fraction (fig. 2). In this case, there seems to be a basis for enrichment of the circumterrestrial swarm in silicate materials by means of the predominant capture of the smaller fraction.

There are currently few experimental data supporting such a conclusion; however, the tendency is completely clear.

Differentiation in composition between the planets and the satellites occurs in the entire solar system but, since the "structural material" is different everywhere, the nature of the fractionation differs for different systems.

Accretion of the Moon and Its Initial Temperature

The obviousness of early heating of the Moon has led many authors to the assumption of its rapid collapse during which a considerable portion of its gravitational energy is retained in the interior. If there was a mechanism which would permit accumulation of the satellite swarm during the entire growth of the Earth (10^8 yr) and then accrete all the material of the swarm into a single mass in a short time ($< 10^3$ yr), the initial temperature of the Moon would have been close to the melting temperature somewhere in the region of the upper mantle (ref. 17). However, this process seems artificial. Assumption of a "short" time scale for growth of the Earth (10^5 yr) cannot help here because the time scale is also too long to retain the energy of accretion.

Let us examine the possibility of accreting the Moon by an acceptable method within the long time scale for growth of the Earth, 10^8 yr. By an acceptable method, we mean satisfaction of the following conditions:

1. Accretion of the Moon must be completed at a distance undoubtedly less than 30 $R_\oplus$ and, probably close to 20 $R_\oplus$, based on the limits given by the tidal evolution of the lunar orbit (ref. 18) and consideration of the geocentric moment of the swarm (ref. 19) ($R_\oplus$ is the radius of the Earth).
2. The initial temperature of the Moon must be higher than some limit $T_0(r)$, in order to ensure heating of the layers responsible for fusion of the continental crust in the "Wasserburg gap" time range (4.6 to 4.0×10^9 yr) (ref. 20), i.e., most likely, within the first 0.3 to 0.4×10^9 yr. Later heating of the deeper layers is responsible for the fusion of the mare basalts. Calculations of Mayeva show that T_0 (r) should

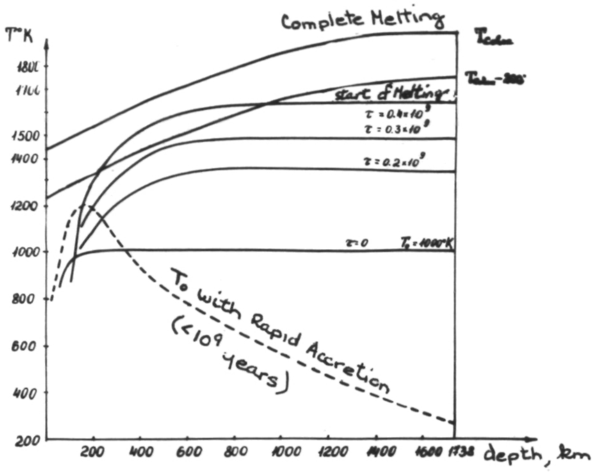

Figure 3.—*Diagram of early thermal history of the Moon. The solid curve was calculated by Maysua for* $U = 7.7 \times 10^{-8}\%$*;* $Th/U = 4$*,* $K/U = 2600$*.*

have the form shown in figure 3, i.e., reach 1000 K or higher in the upper mantle of the Moon and not increase with depth. In particular, the temperature curve for rapid accretion decreases in temperature toward the center (dashed line). The calculation was carried out for a spherical-symmetrical model having a solid outer layer, 700 to 800 km deep at the present time, and a heat flux equal to that observed. However, the initial conditions of the Moon can scarcely be considered to be spherically symmetrical. The present difference in thickness of the continental crust on the visible and the back sides of the Moon and the great irregularities in structure and composition of the Moon indicate the necessity for asymmetry in the construction, composition, and initial temperature distribution. Therefore, still another condition should be imposed.

3. The rather large structural asymmetry of the Moon may be due to the impacts of very large bodies. We note immediately that the probability of collision of the Moon with a large body moving in a heliocentric orbit is very low. Still less likely (ref. 21) is the collision near Earth of two moon-like masses moving in heliocentric orbits.

With these conditions, we will select one of three versions of accretion of the Moon.

a. Formation of an Earth-Moon system from a binary nucleus with initial masses much less than the present ones, mainly by means of absorption of particles from heliocentric orbits
b. Gradual growth of the Moon from a small nucleus, in a swarm where the mass of the Earth was about one-half $n_\oplus$
c. Accretion of the Moon for several large satellites, which grew in the circumterrestrial swarm

Each of these versions can be analyzed in detail, but we note here that the first and second methods lead to a low (about 300 K) initial temperature of the Moon and do not cause any large irregularities in its structure. We shall consider the third version further.

We (ref. 22) analyzed the possibility of forming several large Earth satellites each with its own supply zone. Within 20 to 25 $R_\oplus$, a system of two to three large satellites could have formed with masses ½ to ⅓ $\mu_{☾}$. With tidal friction such a system could have existed only as long as required for close approaches which result in considerable interactive perturbations of the satellites. The rate of tidal removal at such an early epoch is very indefinite. For agreement with the age of the Earth-Moon system, 4.5×10^9 yr, a whole set of functions of the type $\delta_\oplus = \delta_0 + \alpha t^n$ can be used, where $\delta_\oplus$ is the effective lag angle for the entire Earth, t is the time, δ_0 and α are constants, and $n = 1, 2, 3$, etc. (fig. 4). The time for tidal removal of large protomoons from the system to 10 $R_\oplus$ is 10^5 to 10^7 yr and 20 $R_\oplus$ requires up to $10^7 - 5 \times 10^8$ yr, which is 10 to 100 times the removal time of such satellites to $\delta_\oplus = \delta_\oplus$ present $\cong 4°$.

We carried out some numerical experiments for the purpose of determining the evolution of a system of several massive satellites where the satellite orbits were brought together by friction, almost to commensurable 1:1 periods (ref. 23). For simplicity, plane two-satellite systems were studied, with initial circular orbits relative to the Earth. All three bodies were consid-

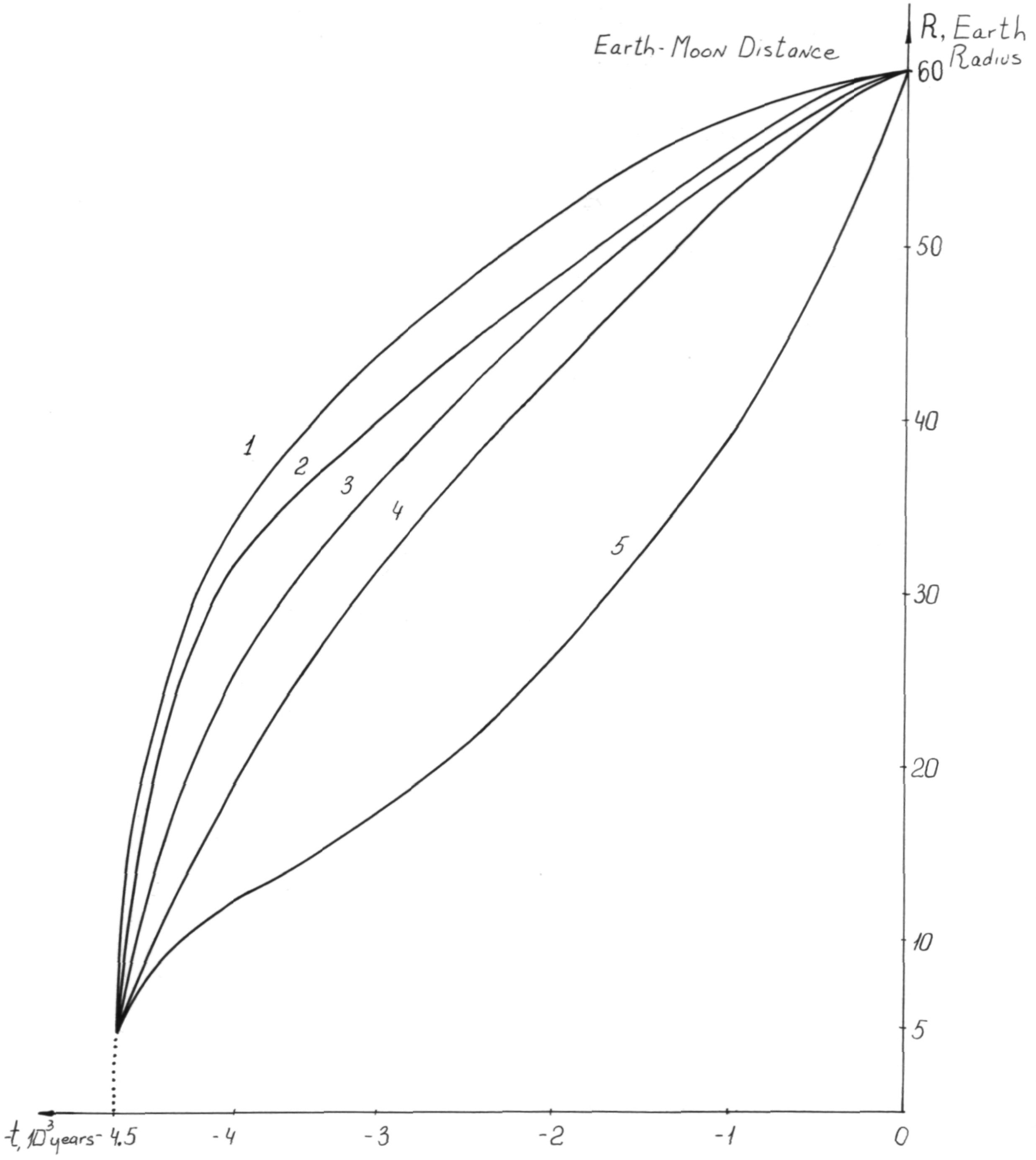

Figure 4.—*Rate of tidal removal of Moon from Earth for various forms of the function* $\delta_{\oplus}$ (p):

1. $\delta \oplus = \delta_0^{I} + \alpha_1 t$;	$\delta \oplus$ *present* $= 1^{\circ}.7$
2. $\delta \oplus = \delta_0^{II}. + \alpha_2 t^2$;	$\delta \oplus$ *present* $= 2^{\circ}.2$
3. $\delta \oplus = \delta_0^{III} + \alpha_3 t^3$;	$\delta \oplus$ *present* $= 4^{\circ}$
4. $\delta \oplus = \delta_0^{IV} + \alpha_4 t^4$;	$\delta \oplus$ *present* $= 4^{\circ}$
5. $\delta \oplus = \delta_0^{V} e^{d_5 t}$	$\delta \oplus$ *present* $= 10^{\circ}$

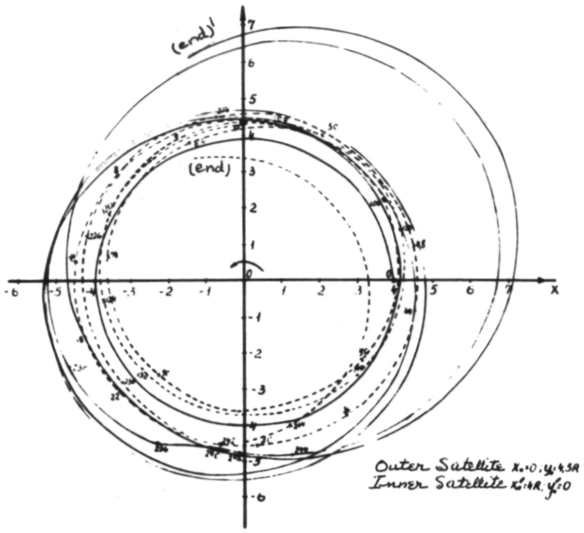

Figure 5.—*Example of the evolution of orbits for two massive satellites, with initial circular orbits at distances of 4 R ⊕ and 4.5 R ⊕ from Earth.*

ered to be physical points. One of the examples of orbital evolution is of two satellites with masses of ½ μ☾ is shown in figure 5. We see strong perturbations of the eccentricities of the orbits, frequent crossing of the orbits, and one case of close passage at a distance of a little less than the sum of the two radii of the satellites, which can be considered to be a collision of the bodies. The rate of relative approach of the satellites is close to parabolic (about 2 km/s); therefore, a collision in systems of this type must always be "frontal," as if the satellites strike one another from an infinitely great distance. The time of evolution of the two-satellite system to a probable collision within 10 $R_⊕$ is numbered in days and, within 20 $R_⊕$, in years. Allowance for three-dimensionality should lengthen the lifetime of the system 10 to 100 yr.

We have not yet carried out such experiments, but, except in the three-dimensional case, the perturbing quantity, except for eccentricity, should be the inclination of the orbit. For such a system, one must consider the precession of the orbits due to the action of the equatorial flattening of the Earth and precession of the axis of the Earth, i.e., a point approximation for the mass of the Earth must be rejected. The problem with two satellites was considered by Goldreich and becomes a much more complicated problem for the Earth-Moon system (ref. 18). One might think that all the limitations indicated by Goldreich for the initial distance of the Moon would be preserved for a system of two protomoons; in particular, in the extreme case one of the two protomoons may form within 10 $R_⊕$ and initially have an "equatorial" orbit. Related to this, we also considered orbits at 4, 6, and 8 $R_⊕$.

A fusion of two satellites into one body is a very rapid process, lasting about 1 hour. The energy given off is sufficient to heat the entire mass to an average temperature several hundred degrees above its equilibrium temperature or to produce a temperature distribution in the interior like that caused by rapid accretion, as mentioned above (see fig. 3). In this manner, formation of the Moon from large bodies, comparable to each other in mass, gives the most acceptable initial temperature from the point of view of subsequent evolution of the Moon. Moreover, the possibility is now open for the creation of large inhomogeneities in structure and composition, as a consequence of the inhomogeneous differentiation of the interior of the Moon.

References

1. SCHMIDT, O. YU., *Four Lectures on Theories of the Origin of the Earth*, USSR Acad. of Sci. Press, Moscow, 1957.
2. SAFRONOV, V. S., *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*, Nauka Press, Moscow, 1969.
3. KAULA, W. M., Dynamical Aspects of Lunar Origin. *Revs. Geophys. Sp. Phys.* Vol. 9, 1971, pp. 217–238.
4. KAULA, W. M. AND A. W. HARRIS, Dynamically Plausible Hypotheses of Lunar Origin, *Nature*, Vol. 245, 1973, pp. 367–369.
5. RUSKOL, YE. L., Cosmogony of the Moon. *Fizika Luny i Planet*, 1972, pp. 160–167.
6. KUIPER, G. P., Comets and the Dissipation of the Solar Nebula. *La Physique des Comets*, 1953, pp. 361–385.
7. RUSKOL, YE. L., Possible Differences in Chemi-

cal Composition of the Earth and Moon in Formation of the Moon in a Circumterrestrial Cluster. *Fizika Luny i Planet*, Vol. 48, No. 6, 1971.
8. GANAPATHY, R., R. R. KEAYS, F. C. LAUL AND E. ANDERS, Trace Elements in Apollo 11 Lunar Rocks. Implications for Meteorite Influx and Origin of the Moon. *Proc. Apollo 11 Lunar Science · Conference, Geochimica y Cosmochimica Acta*, Vol. 2, 1970, pp. 1117–1142.
9. ZVYAGINA, YE. V., G. V. PECHERNIKOVA AND V. S. SAFRONOV, Qualitative Solution of the Coagulation Equation, with Disintegration of Bodies Taken into Account. *Astron. Zhurn*, Vol. 50, 1973, pp. 1261–1273.
10. DOHNANYI, J. F., Fragmentation and Distribution of Asteroids. *Physical Studies of Minor Planets*, 1971, pp. 263–295.
11. RUSKOL, YE. L., Origin of the Moon III. Dynamics of the Circumterrestrial Cluster. *Fizika Luny i Planet*, Vol. 48, No. 4, 1971, pp. 819–829.
12. OROWAN, E., Density of the Moon and Nucleation of Planets. *Nature*, Vol. 222, 1969.
13. DIETZEL, H., G. NEUKUM AND P. RANSER, Micrometeoroid Simulation Studies on Metal Targets. *J. Geophys. Res.*, Vol. 77, No. 8, 1972, p. 1375.
14. KERRIDGE, J. F. AND J. F. VEDDER, Accretionary Processes in the Early Solar System: An Experimental Approach. *Science*, Vol. 177, 1972, pp. 161–163.
15. DREMIN, A. N. AND O. N. BREUSOV, Physical-Chemical Processes in Shock Compression. *Vestn Akad nauk SSSR*, Vol. 9, 1971, pp. 55–59.
16. YUSHKIN, N. P., *Mechanical Properties of Minerals*, 1971.
17. TOKSÖZ, M. N. AND D. H. JOHNSTON, The Evolution of the Moon. *Icarus*, Vol. 21, No. 4, 1974, pp. 389–414.
18. GOLDREICH, P., History of the Lunar Orbit. *Revs. of Geophys.*, Vol. 4, No. 4, 1966, pp. 411–439.
19. RUSKOL, YE. L., Role of the Satellite Cluster in the Origin of Rotation of the Earth. *Astron Vestn*, Vol. 6, No. 2, 1972, pp. 91–95.
20. TERA, F., D. A. PAPANASTASSIOU AND G. J. WASSERBURG, A Lunar Cataclysm at 3.95 AE and the Structure of the Lunar Crust. *Proc. the Fourth Lunar Science Conference*, 1974.
21. RUSKOL, YE. L., Origin of the Moon II. Growth of the Moon in a Circumterrestrial Satellite Cluster. *Astron. Zhurn.*, Vol. 40, 1963, pp. 288–296.
22. RUSKOL, YE. L., Model of Accretion of the Moon, Compatible with Data on Composition and Age of Lunar Rocks. *The Moon*, Vol. 6, 1973, pp. 176–189.
23. RUSKOL, YE. L., YE. V. NIKOLAYEVA AND A. S. SYZDYKOV, Dynamic History of a Plane, Two-Satellite System. *Fizika Luny i Planet*, 1974. In press.